Off-line smoothing

1-sliced

$$p(z_t=k|x_{1:T}) \propto p(z_t=k, x_{1:T})$$

$$= p(x_{1:t-1}, z_t=k)\, p(x_{t:T}|z_t=k)$$

$$= f_t(k)\, p(x_t|z_t=k)\, p(x_{t+1:T}|z_t=k)$$

$$= f_t(k)\, p(x_t|z_t=k)\, b_t(k)$$

2-sliced

$$p(z_t=k, z_{t+1}=k'|x_{1:T}) \propto p(z_t=k, z_{t+1}=k', x_{1:T})$$

$$= p(x_{1:t-1}, z_t=k)\, p(x_t|z_t=k)\, p(z_{t+1}=k', x_{t+1:T}|z_t=k)$$

$$= f_t(k)\, p(x_t|z_t=k)\, p(z_{t+1}=k'|z_t=k)\, b_{t+1}(k')\, p(x_{t+1}|z_{t+1}=k'$$

Sampling states

$$p(z_{1:T+1} | x_{1:T})$$

$$= p(z_1 | x_{1:T})\, p(z_{2:T+1} | z_1, x_{2:T})$$

$$= p(z_1 | x_{1:T})\, p(z_2 | z_1, x_{2:T})\, p(z_{3:T+1} | z_2, x_{3:T})$$

$$= p(z_1 | x_{1:T}) \left[\prod_{t=1}^{T-1} p(z_t | z_{t-1}, x_{t:T}) \right] p(z_{T+1} | z_T)$$

$$p(z_1 = h | x_{1:T}) \propto p(z_1 = h, x_{1:T})$$

$$= p(z_1 = h, x_1)\, p(x_{2:T} | z_1 = h)$$

$$= p(z_1 = h)\, p(x_1 | z_1 = h)\, b_1(h)$$

$$p(z_t = h | z_{t-1}, x_{t:T}) \propto p(z_t = h, x_{t:T} | z_{t-1})$$

$$= p(z_t = h, x_t | z_{t-1})\, p(x_{t+1:T} | z_t = h)$$

$$= p(z_t = h | z_{t-1})\, p(x_t | z_t = h)\, b_t(h)$$

GMM

θ' all param. $\theta'_1, \dots, \theta'_C$ and π'

Complete data $D = (z_1, x_1), \dots, (z_N, x_N)$

$$L(\theta'; D) = \prod_{n=1}^{N} \pi'_{z_n} \, p(x_n \mid \theta'_{z_n})$$

ind. func

$$= \prod_{n} \prod_{c=1}^{C} \left[\pi'_c \, p(x_n \mid \theta'_c) \right]^{I(c = z_n)}$$

$$\ell(\theta'; D) = \sum_{n} \sum_{c} I(c = z_n) \left[\log \pi'_c + \log p(x_n \mid \theta'_c) \right]$$

$$= \sum_{c} \underbrace{\left[\sum_{n} I(c = z_n) \right]}_{N_c} \log \pi'_c + \sum_{c} \underbrace{\sum_{n} I(c = z_n) \log p(x_n \mid \theta'_c)}_{\sum_{n : z_n = c} \log p(x_n \mid \theta'_c)}$$

First term max by $\pi'_c = N_c / N$

EM for GMM

Parameters $\theta = (\pi_1, \dots, \pi_C, \theta_1, \dots, \theta_C)$ where $\theta_c = (\mu_c, \sigma_c^2)$.

Current parameters θ^i.

Expected complete log-likelihood

$$\sum_n E_{p(z_n|x_n,\theta^i)}[\ell(\theta; z_n, x_n)]$$

$$= \sum E\left[\log \prod_c \left(\pi_c\, p(x_n|z_n=c,\theta_c)\right)^{I(z_n=c)}\right]$$

$$= \sum E\left[\sum_c I(z_n=c)\log \pi_c + \sum_c I(z_n=c)\log p(x_n|\theta_c)\right]$$

$$= \sum_{n,c} \underbrace{E[I(z_n=c)]}_{=p(z_n=c|x_n,\theta^i)=r_{nc}} \log \pi_c + \sum_{n,c} \underbrace{E[I(z_n=c)]}_{=r_{nc}} \log p(x_n|\theta_c)$$

$$= \overbrace{\sum_c \left[\sum_n r_{nc}\right] \log \pi_c}^{(A)} + \sum_c \overbrace{\sum_n \underset{n}{r_{nc}} \log p(x_n|\theta_c)}^{(B_c)}$$

(A) and each (B_c) can be maximised separately.

(A) max. by

$$\pi_c = \sum_n r_{nc} / N = \overset{\text{def as } \sum_n r_{nc}}{r_c} / N$$

For (B_c),

$$\text{Let} \quad \alpha_c = \frac{1}{\sigma_c}$$

and

$$f_c(\mu_c, \alpha_c) = \sum_n r_{nc} \log\left[\frac{\alpha_c}{\sqrt{2\pi}} e^{-\frac{\alpha_c^2}{2}(x_n - \mu_c)^2}\right]$$

$$= \sum r_{nc} \log \alpha_c - r_{nc}\frac{\alpha_c^2}{2}(x_n - \mu_c)^2 + C$$

some const.

$$\frac{df}{d\mu_c} = \sum_n r_{nc}\alpha_c^2(x_n - \mu_c) \quad \text{and}$$

$$\frac{df}{d\mu_c} = 0 \Rightarrow \sum_n r_{nc}\cancel{\alpha_c^2} x_n = \sum_n r_{nc}\cancel{\alpha_c^2}\mu_c$$

$$\text{so} \qquad \mu_c = \sum_n x_n r_{nc} \Big/ r_c$$

—

$$\frac{df}{d\alpha} = \sum_n \frac{r_{nc}}{\alpha_c} - r_{nc}\alpha_c(x_n - \mu_c)^2$$

$$\frac{df}{d\alpha} = 0 \Rightarrow \frac{1}{\alpha_c}\sum_n r_{nc} = \alpha_c \sum_n r_{nc}(x_n - \mu_c)^2$$

$$\text{so} \quad \sigma^2 = \frac{1}{\alpha^2} = \sum_n r_{nc}(x_n - \mu_c)^2 \Big/ r_c.$$